

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan penelitian terkait analisis sentimen 5 Destinasi Super Prioritas (DSP) pada media sosial Twitter secara streaming dan real-time menggunakan big data platform Apache Kafka dan Apache Spark menghasilkan kesimpulan sebagai berikut.

1. Penelitian ini berhasil mengembangkan model komputasi untuk melakukan analisis sentimen terhadap Destinasi Super Prioritas pada media sosial Twitter secara streaming dan real-time. Adapun tahapan proses yang dilakukan diantaranya tahap pengumpulan data twitter dengan menggunakan tweet-harvest, tahap streaming dengan Kafka, dan tahap analisis sentimen dan proses analisis lainnya dengan Spark. Kemudian menyajikannya dalam bentuk dashboard menggunakan Dash Plotly.
2. Penelitian ini berhasil mengimplementasikan model komputasi untuk melakukan simulasi analisis sentimen terhadap 5 DSP pada media sosial Twitter secara streaming dengan melakukan beberapa penyesuaian. Dan berhasil menyajikannya hasilnya ke dalam dashboard dengan menampilkan beberapa visualisasi.
3. Hasil eksperimen analisis sentimen secara keseluruhan menunjukkan sebagian besar pengguna Twitter menunjukkan sentimen netral terhadap DSP dengan lebih dari setengah total tweet berada dalam kategori ini yakni sebanyak 61,8%. Sentiment positif sebanyak 27,6% dan sentimen negatif paling sedikit 10,6%. Hasil penilaian *expert judgement* yang menilai berdasarkan konteks kepariwisataan yang dibicarakan pengguna Twitter menunjukkan sebanyak 80% membicarakan konteks pariwisata, sedangkan sebanyak 20% di luar konteks kepariwisataan. Komputasi yang dibuat akan efektif jika sesuai dengan konteks yang akan dianalisis. Twitter cocok digunakan untuk analisis secara streaming dan real time, tetapi tidak terlalu cocok untuk konteks kepariwisataan. Sebagian sentimen di Twitter membicarakan politik, hasil dari 20% yang tidak membicarakan pariwisata sebanyak 60% membicarakan politik. Namun, dari hasil eksperimen yang dilakukan, cukup merepresentasikan hasil yang sesuai dengan aktualnya. Destinasi Borobudur dan Mandalika merupakan destinasi yang paling sering dibicarakan oleh pengguna

Twitter. Sementara itu, destinasi Likupang kurang dikenal atau kurang diperbincangkan, sehingga membutuhkan perhatian lebih dari pemerintah.

4. Penelitian ini menunjukkan bahwa penggunaan Apache Kafka dan Apache Spark efektif dalam mengelola dan menganalisis data Twitter dalam skala besar secara real-time. Performa Spark secara umum efisien, terutama untuk batch yang lebih besar di mana Spark mempertahankan latensi rendah per record meskipun total waktu respons lebih tinggi. Jumlah sentimen yang lebih tinggi tidak berpengaruh dalam meningkatkan latensi per record, menunjukkan bahwa Spark mempertahankan karakteristik skalabilitas yang baik.

5.2. Saran

Penelitian analisis sentimen Destinasi Super Prioritas (DSP) berbasis streaming menggunakan big data platform Apache Kafka dan Apache Spark masih memiliki kekurangan. Salah satunya pada tahap pengambilan data twitter, karena adanya kendala data yang diterima tidak secara langsung masuk ke proses streaming, sehingga diperlukan beberapa penyesuaian. Oleh karena itu, saran-saran yang diberikan dari peneliti terhadap penelitian ini yang dapat digunakan untuk penelitian selanjutnya antara lain sebagai berikut.

1. Tahap pengambilan data twitter agar dapat diproses langsung untuk proses *streaming* dapat dilakukan dengan berlangganan akun Twitter, sehingga dapat menggunakan API Twitter untuk pengambilan data dan dapat secara langsung diproses oleh Kafka untuk proses *streaming*.
2. Dalam proses pengklasifikasian sentimen teks tweet pada Twitter, pendekatan yang digunakan adalah pendekatan berbasis leksikon dengan metode yang digunakan adalah *Dictionary-Based Sentiment Analysis*. Sehingga perlu membuat kamus data positif dan negatif sendiri karena beberapa pustaka untuk pendekatan berbasis leksikon seperti *textblob*, *vader*, dan sebagainya belum memberikan hasil yang baik. Karena proses pengklasifikasian sentimennya dengan menghitung kemunculan kata positif atau negatif dalam kamus data, pendekatan berbasis leksikon memiliki kekurangan dalam menangani konteks dan makna kata. Oleh karena itu, saran untuk penelitian selanjutnya dapat menganalisis berdasarkan konteks dan menerapkan klasifikasi sentimen dengan metode lain seperti menggunakan pendekatan *machine learning* atau *MLLib* pada Spark.

3. Pengembangan dan evaluasi kamus sentimen, jika ingin menggunakan pendekatan *Dictionary-Based Sentiment Analysis* fokus pada pengembangan kamus sentimen khusus yang sesuai dengan konteks destinasi super prioritas. Buat kamus sentimen yang mencakup kata-kata kunci yang paling sering muncul dalam konteks positif, negatif, dan netral terkait destinasi super prioritas. Kamus ini harus terus diperbarui dan diperluas seiring waktu untuk memastikan keterkaitan dan akurasi.
4. Penyimpanan data dalam penelitian ini masih disimpan ke dalam penyimpanan lokal. Untuk penelitian selanjutnya dapat disimpan ke dalam cloud atau pada big data platform lainnya seperti *Hadoop Distributed File System* (HDFS), dan sebagainya.